EARTHQUAKE ENGINEERING AND STRUCTURAL DYNAMICS
Earthquake Engng. Struct. Dyn. **27**, 1465–1481 (1998)

A MEMORY-MATRIX-BASED IDENTIFICATION METHODOLOGY FOR STRUCTURAL AND MECHANICAL SYSTEMS

FIRDAUS E. UDWADIA[1,*] AND WLODEK PROSKUROWSKI[2]

[1]*Department of Civil Engineering, Mechanical Engineering and Business Administration, University of Southern California, OHE 430K, Los Angeles, CA 90089-1453, U.S.A.*
[2]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113, U.S.A.*

SUMMARY

This paper uses an associative memory approach to identify the properties of structural and mechanical systems. The methodology differs from standard identification methods in that it uses a *set* of parameter vectors simultaneously to generate the estimated parameter vector. The method develops a technique for sequentially generating genetically engineered relevant parameter vectors whose use results in accurate identification, while still using small data sets. This makes the approach promising for on-line, rapid, identification of structures and their health monitoring. © 1998 John Wiley & Sons, Ltd.

KEY WORDS: associative memory matrix; rapid parameter identification; genetic iterative algorithms; health monitoring; structural and mechanical systems; linear and non-linear systems

INTRODUCTION

Parameter identification of structural and mechanical systems is an important field of applied mechanics. This is motivated by the need to predict the response of dynamic systems from their models and/or be able to control such systems so that they perform their functions in an adequate and safe manner. Areas such as the active control of structures, the control of robotic manipulators and the non-destructive testing of structures have engendered a host of parameter identification methods.[1–10] However, even for systems which can be reasonably well modelled as being linear, the parameter identification problem often results in a non-linear optimization problem requiring an iterative approach for its solution. Several on-line and off-line methods, many based on the Gauss–Newton approach (or its variants) have been developed. Probabilistic methods such as maximum likelihood estimation and Kalman filtering have also been developed. Yet, when dealing with the identification of a large number of parameters from input-output measurements in complex, spatially extended systems the following two basic difficulties persist: (1) the objective function surface may have multiple maxima and minima so that convergence to the correct parameters is possible only if one starts from a close enough initial guess of the parameters to be identified, and (2) the inverse problem has the inherent possibility of yielding non-unique parameter estimates when using response data obtained from a few locations in a system.

An alternative method which alleviates these difficulties was proposed by Kalaba and Udwadia;[11,12] it uses the conceptual framework of associative memories. For a given input to the system, rather than

* Correspondence to: Firdaus E. Udwadia, Department of Civil Engineering, Mechanical Engineering and Business Administration, University of Southern California, OHE 430K, Los Angeles, CA 90089-1453, U.S.A. E-mail: fudwadia@almaak.usc.edu

CCC 0098–8847/98/121465–17$17·50

Received 15 February 1998

iteratively solving the *inverse* problem starting with a *single* set of parameter values (the so-called initial guess) and successively 'updating' this single set of parameter values, they suggest solving the *forward* problem for *several sets* of parameter values, for the given input. For each set of parameter values the solution of the forward problem yields a corresponding output; thus, the identification scheme they suggest 'associates' to each set of parameter values, its corresponding output. Using a parameter m-vector p_j, they solve the forward problem, for a given input, to obtain the corresponding output n-vector r_j. Thus, the identification scheme relies on providing 'training vectors' (i.e. exposure to various sets of parameter values and the corresponding outputs they produce, for a given input) thereby developing an adequate knowledge base. When the identification scheme is later presented with a given measured response, it can then estimate the parameter set which generated it.

This procedure conceptually departs from the methods usually employed for system identification in three significant respects. Firstly, as mentioned before, instead of using *a single* parameter vector as the initial guess, it *simultaneously* uses several parameter vectors at a time described by the m by N matrix $P = [p_1\ p_2 \cdots p_N]$, and several sets of corresponding outputs described by the n by N matrix $R = [r_1\ r_2 \cdots r_N]$. Secondly, it relies heavily on the solution of forward problems which are usually well posed and yield unique solutions. We assume that the system's model is available and that we can compute the response vector r_j (suitably sampled in time) as a function of the given input, I, to the system, and the parameter vector p_j; hence $r_j = f(p_j; I)$. The usual approaches in system identification rely to a great extent on solving inverse problems which are usually ill-posed. And, thirdly, the aim of this identification approach is to directly build an association between various parameter sets and the corresponding outputs (for a given input). As opposed to this, the usual paradigm employed in system identification, aims to build of an association between various inputs and the corresponding outputs to the system (while starting from a *single* parameter vector, and successively updating it).

Specifically, a simple linear associative m by n memory matrix M defined by the relation $P \approx MR$ is determined by using a suitable cost function and minimizing a suitable matrix norm of the error between P and MR while being attentive to the numerical stability of the results. The association between various parameter sets and the corresponding responses (for a given input) is thus encapsulated in the memory matrix M. Then, when presented with a given measured response r^* of the system, one obtains an estimate, $\hat{p}$, of the parameter vector p^*, as $\hat{p} = Mr^*$.

As pointed out by Kalaba and Udwadia,[11] an important issue for obtaining good results from such an identification scheme concerns the generation of *relevant* training vectors p_j, so that a suitable memory matrix M can be generated. As successively improved training vectors are generated, the memory matrix is successively improved, as is the estimate of the parameter vector. It is this aspect of the identification scheme that this paper principally deals with. We develop here a simple iterative scheme, which is related to genetic algorithms, for the adaptive generation of such training vectors. While comparable in computational intensity to the schemes described in References 11 and 12, this new approach for obtaining relevant training vectors is shown to lead to significant improvements in the capability of the associative memory identification methodology. In Reference 13, Lin and Durand extended the original Kalaba and Udwadia scheme[11,12] employing a weighting algorithm that improves significantly the accuracy of parameter estimation. Our approach contains some elements that are similar to theirs. On the other hand, it has a different emphasis (no weights are used here), with distinctly different results.

THE IDENTIFICATION METHODOLOGY

The identification scheme using memory matrices may be divided into two conceptual steps: (a) determination of the memory matrix M for a given initial set of training vectors $P = [p_1\ p_2 \cdots p_N]$ and a corresponding set of responses $R = [r_1\ r_2 \cdots r_N]$, and (b) generation of *relevant* training vectors. When the system model is

subjected to the input I, each parameter vector p_j when used in the system's model yields a corresponding response vector r_j.

(a) Determination of the memory matrix M. The associative memory approach to system identification relies on finding a memory matrix which minimizes the multicriterion cost function

$$J = (1 - \hat{\alpha})\|MR - P\|^2 + \hat{\alpha}\|M\|^2, \quad 0 \leqslant \hat{\alpha} < 1 \tag{1}$$

where the parameter $\hat{\alpha}$ indicates the relative emphasis placed between the cost associated with mapping the response matrix R to the parameter matrix P (the first term on the right-hand side of equation (1)), and the cost of the computed memory matrix M being ill-conditioned. When the parameter $\hat{\alpha}$ is zero, no penalty is placed on the possible numerical ill-conditioning of the memory matrix M, and all emphasis is placed on obtaining the best mapping; when $\hat{\alpha}$ is close to unity little emphasis is placed on the identification, and much more emphasis is placed on numerical conditioning.

Defining the Euclidean norm of the matrix M as

$$\|M\| = \sqrt{\mathrm{Tr}\{M^{\mathrm{T}}M\}}, \tag{2}$$

the minimization of J with respect to M yields (see Reference 12 for a derivation),

$$M = PR^{\mathrm{T}}\left[RR^{\mathrm{T}} + \left(\frac{\hat{\alpha}}{1 - \hat{\alpha}}\right)I\right]^{-1} = PR^{\mathrm{T}}[RR^{\mathrm{T}} + \alpha I]^{-1} \tag{3}$$

where we have denoted, for convenience, the ratio $(\hat{\alpha}/(1 - \hat{\alpha}))$ by α. It should be noted that while the result obtained in equation (3) is similar to the least-squares result with the additional use of the regularization parameter α, we now have a greater physical understanding of the role played by this regularization parameter α in its effort to alter the cost function J from one which emphasizes accurate identification at the expense of possible numerical ill-conditioning (α, a positive number close to zero) to one which emphasizes numerical conditioning at the expense of accurate identification (α, a large positive number).

(b) Generation of relevant training vectors. The use of a set of parameter vectors to determine the linear mapping $M : R \to P$ works well even when the initially guessed parameter vectors are 'far' from the actual parameter vector which describes the system. It was pointed out in Reference 12 that the methodology would stand to considerable gain if a sequential procedure for determining additional *relevant* training vectors could be devised so that by using these training vectors, sequential updates of the estimate of the parameter vector could be obtained. Such an updating scheme was developed; to the prevalent pool of training vectors, at each iteration was added a randomly generated training vector which was 'close' to the current parameter estimate. Thus, the pool of training vectors was expanded gradually by one training vector at each iteration, see Reference 12 for details. In this paper we present a superior way of generating training vectors, which we have found to be efficacious in handling a number of difficult identification problems.

Consider a dynamic system which is subjected to an input which may or may not be time-varying, and whose response r^* to this input is measured. For example, we could have a five-storey building structure modelled as a five-degree-of-freedom system subjected to an (input) impulsive force, or to strong earthquake ground shaking at its base. Its response can be measured at, say, the first- and second-storey levels over certain time windows and the response data (say equally spaced in time) concatenated into an n-vector r^*. If the stiffness distribution along the structure is to be identified, the true parameter vector p^* would contain the five stiffnesses at each of the five storey levels. The aim is then to identify the 5-vector p^* from the input data and the measured n-vector of response, r^*.

The iterative identification scheme we develop in this paper can be schematically described as follows:

1.0 *Initialization*:

(a) Starting with a set of initial guesses (initial training vectors) $P^{(0)} = [p_1^{(0)}\ p_2^{(0)} \cdots p_{N_0}^{(0)}]$ of the true parameter vector p^*, generate a set of corresponding responses $R^{(0)} = [r_1^{(0)}\ r_2^{(0)} \cdots r_{N_0}^{(0)}]$, when the system model is subjected to the given specific input, I; the initial training vectors, $p_j^{(0)}$, are selected using our physical understanding of the system and our knowledge of the ranges in which the parameter values may lie. The response to the specific input (suitably sampled at certain intervals of time, say) is obtained from our knowledge of the 'structure' of the system model, and can be expressed as $r_j^{(0)} = f(p_j^{(0)}; I)$.
(b) Compute the initial memory matrix $M^{(0)}$ using equation (3), with a suitable α.
(c) Using the measured response, r^*, of the actual system when subjected to the same input I, determine a primary estimate of the true parameter vector as $\hat{p}^{(0)} = M^{(0)} r^*$.

2.0 *Iterate for* $k = 1, 2, \cdots$ *until convergence*

(a) Based on the estimate $\hat{p}^{(k-1)}$, generate a *relevant* set of training vectors $p^{(k)} = [p_1^{(k)}\ p_2^{(k)} \cdots p_{N_k}^{(k)}]$ and compute on the basis of the model, the corresponding responses $R^{(k)} = [r_1^{(k)}\ r_2^{(k)} \cdots r_{N_k}^{(k)}]$ to the given input I.
(b) Using equation (3), determine the updated memory matrix $M^{(k)}$, as $M^{(k)} = P^{(k)} R^{(k)\mathrm{T}} [R^{(k)} R^{(k)\mathrm{T}} + \alpha_k I]^{-1}$, with a suitable α_k.
(c) Compute the updated parameter estimate $\hat{p}^{(k)} = M^{(k)} r^*$.

As stated earlier, the initialization step is primarily concerned with the development of the primary memory matrix $M^{(0)}$ and a primary estimate, $\hat{p}^{(0)}$, of the parameter vector; the second step deals with the creation of a relevant pool of training vectors based on this initial estimate $\hat{p}^{(0)}$ so that one iteratively generates the updated memory matrices $M^{(k)}$, $k > 0$, and thereby iteratively obtains improved estimates of the parameter vector.

We note that despite the different numbers, N_k, of training vectors that can be used at each iteration, k, the updated estimate, $\hat{p}^{(k)}$, of the parameter vector is a single vector. In this paper we start with five initial training vectors (i.e. $N_0 = 5$) and increase their number to $N_k \equiv \bar{N}$ for all future iterations, $k > 0$. Also, we use in our procedure a constant value of α_k, with $\alpha_k \equiv \alpha$, $k \geqslant 0$.

As seen from equation (3), the matrix $M^{(k)}$ depends on the parameter α. The larger α is, the more emphasis, in general, is given to the identification of the computed parameter vector, the less to stability of the memory matrix. Consequently, one should keep α as small as possible, but large enough so that $M^{(k)}$ is sufficiently well-conditioned. The poor conditioning of the memory matrix when using noisy data may lead one to increase the value of α. In the examples considered in this paper, the value of α was kept between 10^{-6} and 10^{-11}.

One way of obtaining the new training set $P^{(k)}$ at the kth iteration from the single estimate $\hat{p}^{(k-1)}$ is to randomly perturb this parameter estimate vector by a small amount β_k. Thus one obtains

$$p_{j,l}^{(k)} = (1 + \theta_{j,l}\beta_k)\hat{p}_j^{(k-1)} \quad \text{for } l = 1, 2, \ldots, \bar{N};\ j = 1, 2, \ldots, m \tag{4}$$

Here we denote by $p_{j,l}^{(k)}$ the jth component of the m-vector $p_l^{(k)}$, and by $\theta_{j,l}^{(k)}$ a random number uniformly distributed in $(-0{\cdot}5, 0{\cdot}5)$. It should be noted that each of training vectors in the set $P^{(k)}$ are determined *anew* at each iteration. This idea is based on the assumption that at each new iteration we obtain a better approximation to the true parameter vector. Contrary to what was done in Reference 12, here new training vectors are *not* appended at each iteration to the set of old ones thereby expanding the gene pool. In order to cluster the new training vectors around the true parameter vector, one may decrease this perturbation β_k as the iterations progress. In what follows, we provide a way of selecting β_k adaptively, based on the normalized response residual.

Let us denote the error in the parameter estimate at the kth iteration by $e^{(k)} = p^* - \hat{p}^{(k)}$; the response residual (for the given input I) corresponding to $\hat{p}^{(k)}$ can accordingly be denoted by $r_E^{(k)} = r^* - \hat{r}^{(k)}$, where, $\hat{r}^{(k)} = f(\hat{p}^{(k)}, I)$. The first quantity, $e^{(k)}$, is not computable (except in test cases) since we have no knowledge of the true parameter p^*, while the second one, can be computed since the true (measured) response of the system r^* corresponding to the true parameter p^* is known, and the system model is also assumed known. The parameter β_k can now be chosen to be a function of the norm of the residual $r_E^{(k)}$. For $k > 0$, a simple choice such as $\beta_k = \rho_k \|r_E^{(k)}\|$ appears to be sufficient (see Figure 2(a)), where the ρ_k's are constants. Here the norm is defined as $\|r\| = \sqrt{(1/n)\sum_{i=1}^{n} r_i^2}$ in order to make β independent of n, the length of the vector r. More generally, one may view the choice of β_k as indicating how changes in the parameter space affect the response space. Using a local linearization, we can derive an approximate proportionality relation

$$\beta_k = \beta_{k-1} \frac{\|r_E^{(k)}\|}{\|r_E^{(k-1)}\|}, \quad k = 1, 2, \cdots \tag{5}$$

which leads to

$$\beta_k = \frac{\beta_0}{\|r_E^{(0)}\|} \|r_E^{(k)}\| = \rho_k \|r_E^{(k)}\|, \quad k = 1, 2, \ldots \tag{6}$$

In the algorithm used here, we have taken $\rho_k \equiv 1/\|r^*\|$, $k > 0$.

The choice of relevant training vectors discussed above can be significantly enhanced from ideas related to the development of gene pools. At each iteration one can choose from among the randomly generated parameter vectors provided by relation (4) those vectors which lead to small residuals, thereby creating a 'good genetic pool' of relevant training vectors. To do this we generate, at each iteration, as many candidates as needed to obtain $\bar{N}$ relevant training vectors $p_i^{(k)}$, let us call them mutations. A mutant generated the kth iteration is added to the gene pool of relevant training vectors if the residual of the response it generates is less than a threshold value. A natural choice, for $k > 0$, is the norm of the residual computed at the previous iteration, normalized with respect to the measured response, namely $\delta_k \|r_E^{(k-1)}\|$. In this paper we choose $\delta_k \equiv \bar{\delta} = 1{\cdot}5$, $k > 0$. Smaller values of $\bar{\delta}$ are found to lead to an excessive number of mutations which do not meet the pool criteria and therefore have to be rejected as genetically inferior; larger values of $\bar{\delta}$ deteriorate the quality of the gene pool. Thus, at each iteration, k, a new gene pool of $\bar{N}$ training vectors is created.

The final form that our genetic iterative algorithm then takes is:

1.1. Initialization
- (a) based on a physical understanding of the system, determine $\hat{p}^{(0)}$ as before in **1.0.**
- (b) compute $\hat{r}^{(0)}$ and $\|r_E^{(0)}\|/\|r^*\|$.
- (c) set $\beta_1 = \rho \|r_E^{(0)}\|$, $\rho = 1/\|r^*\|$.

2.1. Iterate for $k = 1, 2, \ldots$ until $\|r_E^{(k)}\| < \tau$.
- (a) set δ_k, $k > 0$.
- (b) generate mutations (candidate training vectors)
  - (i) $p_{j,l}^{(k)} = (1 + \theta_{j,l}\beta_k)\hat{p}_j^{(k-1)}$, for $l = 1, 2, \ldots; j = 1, 2, \ldots, m$.
  - (ii) compute the responses $r_l^{(k)} = f(p_l^{(k)}; I)$ based on the system model for the given input, I.
  - (iii) accept $\bar{N}$ vectors $p_l^{(k)}$ which satisfy the criterion $\|r^* - r_l^{(k)}\| < \delta_k \|r_E^{(k-1)}\|$.
- (c) form (i) the set $P^{(k)} = [p_1^{(k)}\ p_2^{(k)} \ldots p_{\bar{N}}^{(k)}]$.
  - (ii) the set $R^{(k)} = [r_1^{(k)}\ r_2^{(k)} \ldots r_{\bar{N}}^{(k)}]$.
- (d) compute the memory matrix $M^{(k)} = P^{(k)} R^{(k)\mathrm{T}} (R^{(k)} R^{(k)\mathrm{T}} + \alpha I)^{-1}$.
- (e) compute the
  - (i) parameter estimate $\hat{p}^{(k)} = M^{(k)} r^*$.
  - (ii) estimated response $\hat{r}^{k} = f(\hat{p}^{(k)}; I)$, and the residual $r_E^{(k)} = r^* - \hat{r}^{(k)}$.
  - (iii) perturbation $\beta_{k+1} = \rho \|r_E^{(k)}\|$.

Preferably, the iterations should be terminated by monitoring the history of the convergence of each of the components of the parameter vector, i.e. $\|\hat{p}_j^{(k)} - \hat{p}_j^{(k-1)}\|$, $j = 1, 2, \ldots, m$. In this paper the identification algorithm was halted after a fixed number (usually 10) of iterations, mostly to obtain a uniform display of convergence history.

The computational cost of this algorithm is dominated by the cost of computing the response of the system to the mutant training vectors. A much smaller contribution to the cost is the generation of the successive memory matrices and the resulting parameter estimates. In Reference 12 the number of training vectors was initially smaller, but grew linearly with the number of iterations; in the approach presented here, convergence occurs with a much smaller number of iterations though the number of training vectors is larger. Consequently, the computational cost of both algorithms is comparable, but the accuracy of the results presented herein is far superior.

We now present two examples of the ability of the above-mentioned identification scheme to handle difficult inverse problem.

NUMERICAL RESULTS

It was shown in Reference 12 that the associative memory approach to system identification may be in certain cases superior to schemes such as the recursive prediction error scheme. Here we consider two examples of identification problems which have not responded well to methods like Kalman filtering and recursive-error-prediction methods. The numerical experiments were done using MATLAB with a machine precision of about $2{\cdot}2 \times 10^{-16}$. In all the experiments, integration was performed using the fourth-order Runge-Kutta method, and the fixed time step chosen for integration was $\Delta t = 0{\cdot}02$.

Example 1. (Five degree-of-freedom building structure). A five-degree-of-freedom building structure (see Figure 1) represented by the equation

$$M\ddot{x} + C\dot{x} + Kx = f(t), \quad x(0) = x_0, \ \dot{x}(0) = \dot{x}_0 \tag{7}$$

is considered where the 5 by 5 matrix M is taken to be the identity matrix. The matrix K is tridiagonal, and has the form

$$K = \begin{bmatrix} k_5 & -k_5 & & & \\ -k_5 & k_4 + k_5 & -k_4 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \\ & & & -k_2 & k_1 + k_2 \end{bmatrix} \tag{8}$$

The matrix C is also tridiagonal and has the a similar form (with damping parameters c_i's replacing the k_i's). The aim is to identify the 10 physical parameters which describe the system $k_j, c_j, \ j = 1, 2, \ldots, 5$. The exact parameters are taken to be

$$\begin{array}{lllll} k_1 = 2080, & k_2 = 2050, & k_3 = 2020, & k_4 = 2100, & k_5 = 2005 \\ c_1 = 1{\cdot}5, & c_2 = 1{\cdot}8, & c_3 = 2, & c_4 = 2{\cdot}5, & c_5 = 2 \end{array}$$

and we denote the parameter vectors $p_s = [k_1\ k_2\ k_3\ k_4\ k_5]$, and $p_c = [c_1\ c_2\ c_3\ c_4\ c_5]$.

The system is not classically damped. All the parameters are taken to be in consistent units.

A. Response to initial velocity. We subject the system to an initial velocity condition expressed by the vector $\dot{x}(0) = [\dot{x}_1(0)\ \dot{x}_2(0)\ \dot{x}_3(0)\ \dot{x}_4(0)\ \dot{x}_5(0)]^{\mathrm{T}} = [10\ -10\ 10\ -10\ 0]^{\mathrm{T}}$. The initial displacement vector $x(0)$ is zero, as is the vector $f(t)$. Table I shows the five initial training vectors used in the identification and the

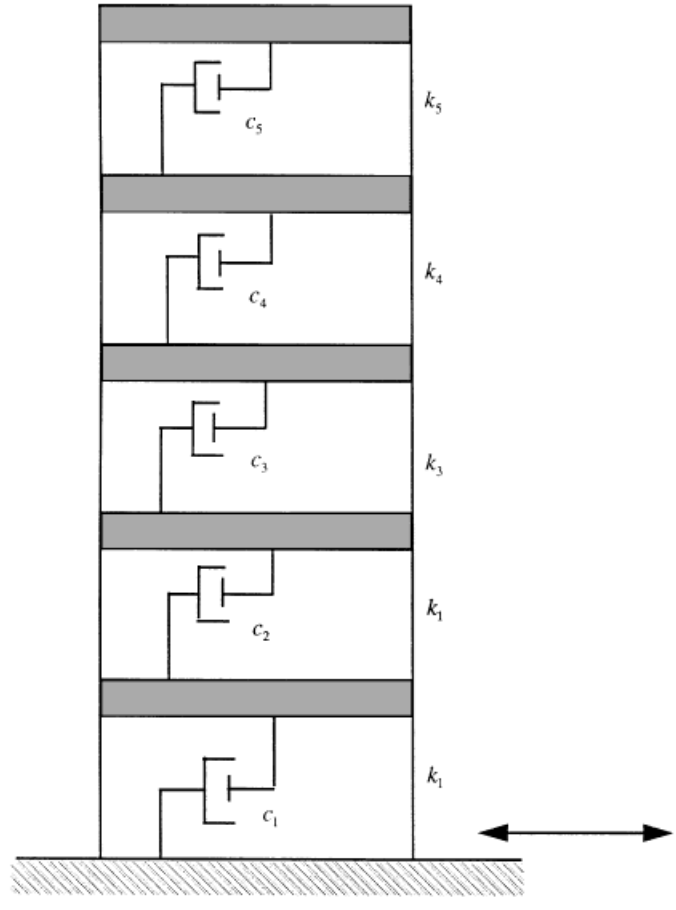

Figure 1. Five storey building subjected to: (a) initial velocity; (b) El Centro 1940, S00W

Table I. Five dof building system: errors in initial training vectors

	Five training vectors, $P^{(0)}$					Errors in $P^{(0)}$(%)					RMS error %
	p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5	
c_5	2·5	4·0	3·0	2·5	3·0	25·0	100·0	50·0	25·0	50·0	57·0
c_4	2·0	3·0	2·0	2·0	3·0	− 20·0	20·0	− 20·0	− 20·0	20·0	20·0
c_3	4·0	2·7	2·8	2·5	3·0	100·0	35·0	40·0	25·0	50·0	56·5
c_2	2·0	2·0	2·0	2·0	3·0	11·1	11·1	11·1	11·1	66·6	31·4
c_1	2·5	1·9	1·0	2·5	3·0	66·7	26·7	− 33·3	66·7	100·0	64·4
k_5	1950	2050	1960	1900	2100	− 2·74	2·24	− 2·24	− 5·24	4·74	3·67
k_4	1900	2100	2000	1950	2100	− 9·52	0·0	− 4·76	− 7·14	0·0	5·73
k_3	2100	1900	2050	2100	2100	3·96	− 5·94	1·48	3·96	3·96	4·11
k_2	2000	2020	2200	2200	2100	− 2·44	− 1·46	7·32	7·32	2·44	4·92
k_1	2200	1940	1900	1950	2100	5·77	− 6·73	− 8·65	− 6·25	0·96	6·22

Table II. Five dof building system: measurements at first and second floors, $\alpha = 10^{-10}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(10)}$		Exact param.
	Error %	Param.	Error %	Param.	
c_5	− 69·59	3·39	0·0381	1·9992	2·00
c_4	− 9·67	2·74	− 0·0106	2·5003	2·50
c_3	− 100·01	4·00	0·0053	1·9999	2·00
c_2	− 40·26	2·52	− 0·0110	1·8002	1·80
c_1	− 29·80	1·95	− 0·0302	1·5005	1·50
k_5	− 7·6	2156·8	− 0·001	2005·02	2005
k_4	− 0·7	2115·3	− 0·001	2100·03	2100
k_3	− 4·5	2111·1	− 0·001	2020·02	2020
k_2	− 0·8	2067·0	− 0·002	2050·04	2050
k_1	− 7·2	2229·0	− 0·002	2080·04	2080

percentage errors in each of the initial guesses of the different parameters. The last column in Table I shows the root-mean-square-error in each parameter value across the set of 5 initial parameter vectors, and equals

$$\sqrt{\frac{1}{5}\sum_{i=1}^{5}\left[\frac{\hat{p}_{ij}-p_i^*}{p_i^*}\right]^2}, \quad j = 1, 2, \ldots, 10.$$

We have assumed that our initial estimates of the stiffness parameters are better than those of our damping parameters, a situation common in the description of building structures. Errors (RMS) in the initial damping estimates are approximately 20–65 per cent, those in the stiffness estimates are approximately 4–6 per cent.

The equations are integrated using the fourth-order Runge–Kutta method and only the responses (displacements) at the first and second storeys are used for the identification. Only in Table III do we use responses from the fourth and fifth storeys for the purpose of comparison. Furthermore, the displacement data from two short time windows of response, $t \in [0, 0{\cdot}6]$ and $t \in [0{\cdot}7, 2{\cdot}5]$, are taken for the identification procedure. The data are concatenated to form the response vector r_j. At each iteration, a gene pool of $\bar{N} = 51$ training vectors is created. Of the mutants generated at each iteration only those that satisfy the criterion described in Step 2.1(b) (iii) of the method (with $\bar{\delta} = 1{\cdot}5$) are included in the gene pool. This leads to a rejection, on average, of about one mutant or less out of every two which are produced. This value of $\bar{\delta}$ was kept fixed for all the examples shown in this paper. The regularization parameter α is chosen to be 10^{-10} (unless indicated otherwise), and the algorithm is halted after $k = 10$ iterations.

Table II shows the primary estimate $\hat{p}^{(0)}$ obtained after the 5 initial parameter vectors are used in the identification process, and the final estimate, $\hat{p}^{(10)}$, after 10 iterations. Also shown are the percentage errors in each of the parameters. The convergence history plotted in Figure 2(a) shows the normalized parameter error $\|\hat{p}_S^{(k)} - p_S^*\|/\|p_S^*\|$ in stiffness and $\|\hat{p}_c^{(k)} - p_c^*\|/\|p_c^*\|$ in damping (in per cent) as a function of the number of iterations, k. Also shown is the normalized residual error $\|r^* - r^{(k)}\|/\|r^*\|$ between the response using the actual and the estimated parameters. As seen from the last two columns of Table II, the procedure yields *very* accurate estimates of the physical parameters. The reduction in the errors in damping and stiffness values is by about 3 orders of magnitude. We note that the normalized residual error curve closely follows the curve for error in stiffness, indicating that the residual is more so dominated by errors in the stiffness, than by those in damping. It should also be noted that using the norm of the difference of consecutive parameter estimates, $\|\hat{p}^{(k)} - \hat{p}^{(k-1)}\|$, as the stopping criterion, the iterations could have been terminated after only 4 iterations with considerable computational savings. We have chosen a fixed number of 10 iterations for a more uniform display, and as a verification that the algorithm is robust and does not diverge as the iteration number

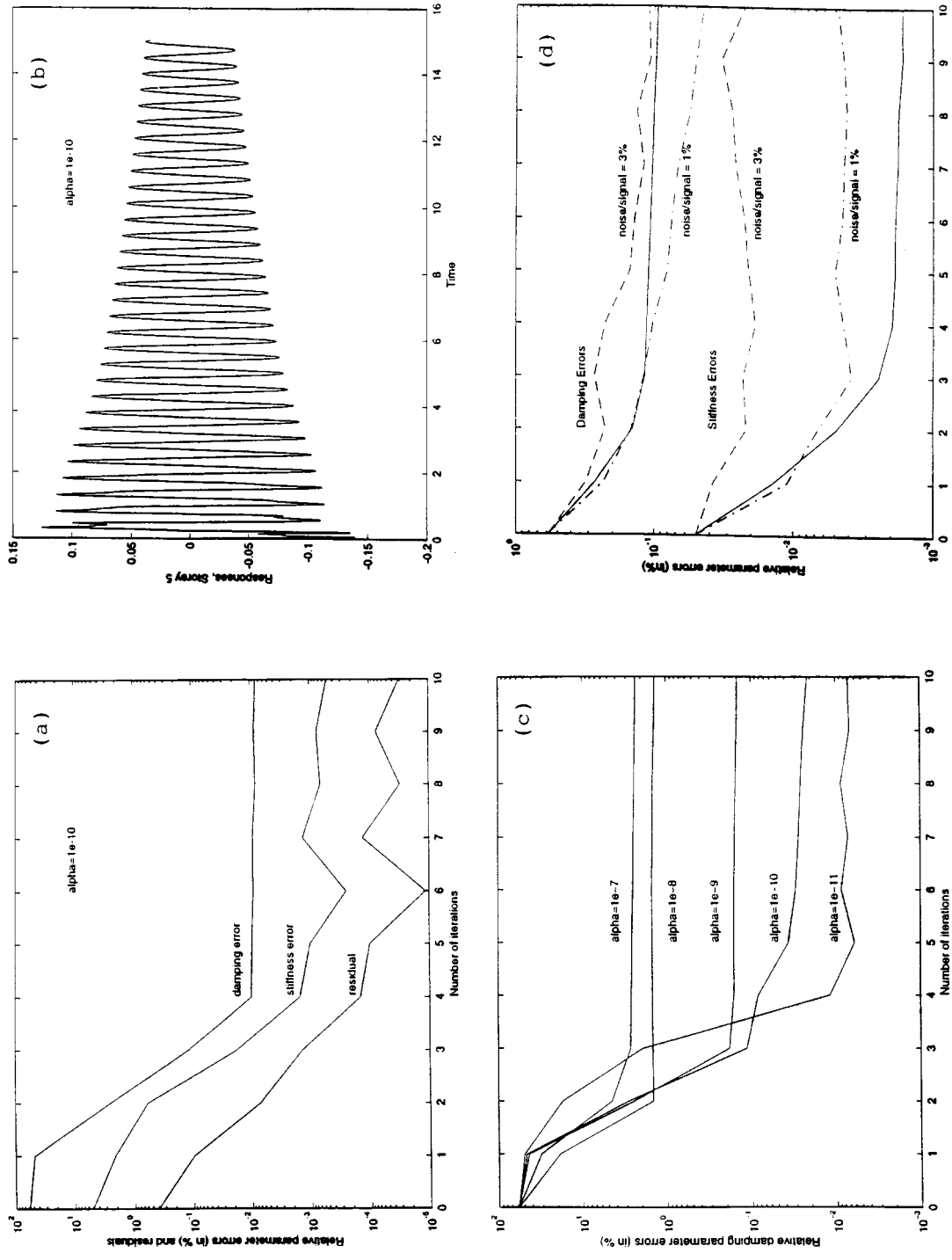

Figure 2. Response to initial velocity

Table III. Five dof building system: Measurements at fourth and fifth floors $\alpha = 10^{-10}$

	Primary estimates. $\hat{p}^{(0)}$		Final estimates. $\hat{p}^{(10)}$		Exact param.
	Error %	Param.	Error %	Param.	
c_5	− 76·18	3·52	− 0·0006	2·00001	2·00
c_4	− 5·40	2·63	− 0·00002	2·5000005	2·50
c_3	− 100·47	4·01	− 0·0048	2·0001	2·00
c_2	− 45·64	2·62	0·0019	1·79997	1·80
c_1	40·57	0·89	0·0039	1·49994	1·50
k_5	− 8·3	2171·6	0·001	2004·98	2005
k_4	− 1·5	2130·7	0·001	2099·97	2100
k_3	− 3·5	2090·5	− 0·0001	2020·003	2020
k_2	− 2·4	2100·2	− 0·002	2050·04	2050
k_1	− 3·4	2151·1	− 0·002	2080·04	2080

increases. Figure 2(b) shows the responses measured at the fifth-storey level and those obtained using the estimated parameters, for the results depicted in Figure 2(a). To the scale of the figure, the difference is indistinguishable.

Table III shows similar results were we to use displacement measurements from only the fourth and fifth storeys over the same two time windows for the identification, all other things remaining unchanged. Again, the last two columns of the table show that the identification of the physical parameters is *very* good. Comparing the results in Tables II and III, we observe that for this input (and the measurement time windows chosen) the identification results obtained, especially for the damping estimates, are superior when using responses from the two upper storeys rather than the two lower storeys of the structure, though both sets of data yield excellent parameter estimates.

The parameter values to be identified in our structural system differ by three orders of magnitude. It is this disparity in the magnitude of the numbers to be identified that causes the identification problem to become difficult; if one were using standard optimization methods, because such techniques would lead to climbing down narrow valleys, the accuracy of the estimated damping parameters would be poor. Our genetic iterative scheme using the memory matrix approach appears to have little difficulty (see, Tables II and III), and in fact usually converges to the final estimates in less than five iterations (see Figures 2(a) and 2(c)) while using data over small intervals of time. This makes the identification approach described herein a good candidate for real time on-line structural identification and health monitoring.

The regularization parameter α was found to have a significant effect on the ability of the identification scheme to estimate the parameters accurately, especially the small numbers describing the damping values. Figure 2(c) shows the effect of α on the accuracy of the identified damping parameters, c_i when using displacement data from the first and second storeys. As expected the smaller α is, the greater is the emphasis placed on identification as opposed to numerical stability, and the accuracy of the identification improves. The figure indicates that a drop in α by an order of magnitude, improves the accuracy of the damping estimate by about the same amount. The optimal value of α depends also on the computer machine precision, here $\approx 2{\cdot}2 \times 10^{-16}$.

Table IV shows the results obtained when noisy measurements are used. The displacement measurements from the first and second storeys are corrupted using a uniformly distributed random noise with a noise-to-signal ratio of 3 per cent. The same initial training vectors were used. We see that the estimates of the parameters deteriorate substantially, especially the damping estimates. The measurement noise is enough to mask the effect of small differences in the damping parameter values, and the identification scheme performs poorly when trying to estimate them. Due to greater inconsistency in the equation $P = MR$ in the presence of measurement noise, a greater value of α is needed to ensure numerical stability.

Table IV. Five dof building system: noise/signal = 3%, $\alpha = 10^{-6}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(10)}$		Exact param.
	Error %	Param.	Error %	Param.	
c_5	− 67·99	3·36	19·7650	1·6047	2·00
c_4	− 9·01	2·73	14·7617	2·1310	2·50
c_3	− 101·29	4·03	− 4·4948	2·0899	2·00
c_2	− 38·62	2·49	− 9·7925	1·9763	1·80
c_1	− 34·86	2·02	− 21·3471	1·8202	1·50
k_5	− 7·1	2148	− 0·7	1991	2005
k_4	− 0·2	2104	− 3·0	2162	2100
k_3	− 4·3	2108	− 0·3	2026	2020
k_2	− 0·2	2055	− 2·3	2096	2050
k_1	− 7·4	2235	3·4	2150	2080

Table V. Five dof building system: error RMS norms in final estimates

	Error norms in final estimates			
	Without noise		With 1% noise	
	$\alpha = 10^{-9}$	$\alpha = 10^{-10}$	$\alpha = 10^{-6}$	$\alpha = 10^{-7}$
Damping	0·145%	0·021%	4·18%	3·73%
Stiffness	0·004%	0·002%	0·41%	0·77%

Figure 2(d) shows the norm errors in stiffness and damping as a function of iteration number for various noise levels, using displacement responses from the first and second storeys with $\alpha = 10^{-6}$, all the other parameters being those used earlier. Table V summarizes the effect of using different α values and the effect of noise on the identification results. It should be noted that the identification even in the presence of measurement noise yields acceptable results, especially when one considers that these results are obtained by looking at very short response time windows for the identification.

B. Response to earthquake excitation. The identification procedure is next applied using the response of the system to earthquake excitation. The structure is subjected to the base acceleration recorded for the El Centro S00W component, and response data from a time windows $t \in [0, 5]$ is used for the identification. Using the same five initial parameter guess vectors as in Table I, and response data only from the first and second storeys, the results of the identification are shown in Table VI. Figure 3(a) shows the errors as a function of the iteration number and Figures 3(b)–3(d) show the measured and estimated responses for $t \in [0, 15]$ at the fifth, fourth and third storey levels. At the scale of the figure, no discrepancies between the two sets of responses are observable. Despite this, Table VI shows that the first three damping parameters are poorly identified indicating that little information about these parameters is available from the input and the response data used in the identification.

Table VII shows results for the identification procedure had we measured displacement responses from the fourth and fifth storeys, and only used these in the identification. Again the data for the identification is gathered over the same time window. Figure 4 shows these results. Table VIII summarizes these results and we observe that for this particular input acceleration and measurement time window chosen, the identification results obtained using data from the lowest two storeys are superior to those obtained from data gathered at the higher two storeys.

Table VI. Earthquake, measurement at first and second floors: $\alpha = 10^{-11}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(10)}$		Exact param.
	Error %	Param.	Error %	Param.	
c_5	− 40·36	2·81	− 35·89	2·718	2·00
c_4	26·64	1·83	22·99	1·925	2·50
c_3	− 107·98	4·16	− 9·76	2·195	2·00
c_2	− 13·99	2·05	1·27	1·777	1·80
c_1	38·97	0·92	− 0·12	1·502	1·50
k_5	− 0·4	2012	0·3	1999·4	2005
k_4	6·5	1963	0·01	2100·1	2100
k_3	− 4·6	2144	− 0·1	2022·2	2020
k_2	− 2·1	2093	0·1	2047·1	2050
k_1	− 3·0	2141	0·1	2077·2	2080

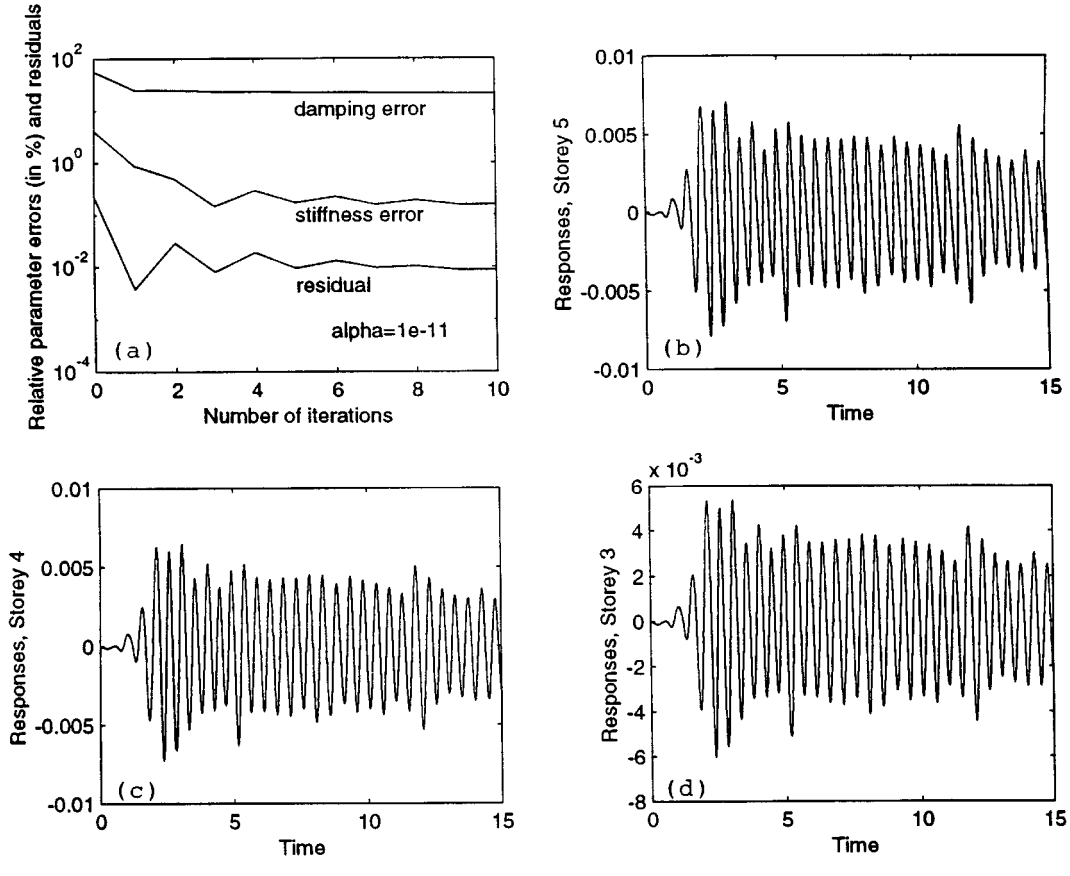

Figure 3. Response to earthquake excitation

Table VII. Earthquake, measurement at fourth and fifth floors: $\alpha = 10^{-11}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(10)}$		Exact param.
	Error %	Param.	Error %	Param.	
c_5	− 27·62	2·55	− 1·08	2·022	2·00
c_4	47·24	1·32	2·74	2·431	2·50
c_3	− 179·36	5·59	− 95·45	3·909	2·00
c_2	19·40	1·45	33·39	1·199	1·80
c_1	51·81	0·72	43·19	1·852	1·50
k_5	3·5	1934	0·02	2004·7	2005
k_4	14·5	1796	1·0	2079·1	2100
k_3	− 2·5	2071	− 0·4	2027·9	2020
k_2	9·0	1866	1·5	2020·2	2050
k_1	11·9	2327	− 1·3	2107·0	2080

Table VIII. Earthquake: error RMS norms in estimates

	1st–2nd floors		4th–5th floors	
	primary	final	primary	final
Damping	55·85%	21·17%	88·04%	47·38%
Stiffness	3·98%	0·16%	9·62%	1·00%

Example 2. (Coupled non-linear oscillators). Consider the coupled Van-der-Pol-Duffing oscillator described by the equations

$$\ddot{x} + a_1(b_1 - x^2)\dot{x} + c_1(x - y) + d_1x^3 = -e\sin(\omega t) \tag{9a}$$

$$\ddot{y} + a_2(b_2 - x^2)\dot{x} - c_2(x - y) + d_2y^3 = 0{\cdot}2\cos(3t) - 0{\cdot}5\cos(2t) \tag{9b}$$

with

$$x(t = 0) = 0,\ \dot{x}(t = 0) = 0,\ y(t = 0) = 0{\cdot}2,\ \dot{y}(t = 0) = 0{\cdot}1 \tag{9c}$$

where the dots represent differentiation with respect to time, t. The exact values of the constant parameters are taken to be

$$a_1 = 5,\quad b_1 = 2,\quad c_1 = 5,\quad d_1 = 2,\ e = 1$$

$$a_2 = 4,\quad b_2 = 2{\cdot}5,\ c_2 = 5,\quad d_2 = 3,\ \omega = 2$$

As seen from the parameter values, the system is highly non-linear; such systems pose substantial difficulties for many identification schemes (see Reference 12). All ten parameters are to be identified using the methodology presented in the previous section. We note that the identification includes both the system parameters *and* parameters related to the forcing function.

Five initial parameter vectors shown in Table IX form the set $P^{(0)}$. Also shown are the percentage errors in each parameter relative to the exact parameter values. We notice that the initial guesses differ substantially from the exact values, and the *qualitative* behavior of the coupled system is strongly dependent on the parameter values. As before, the RMS errors in each of the estimates are also shown. Two short disjoint

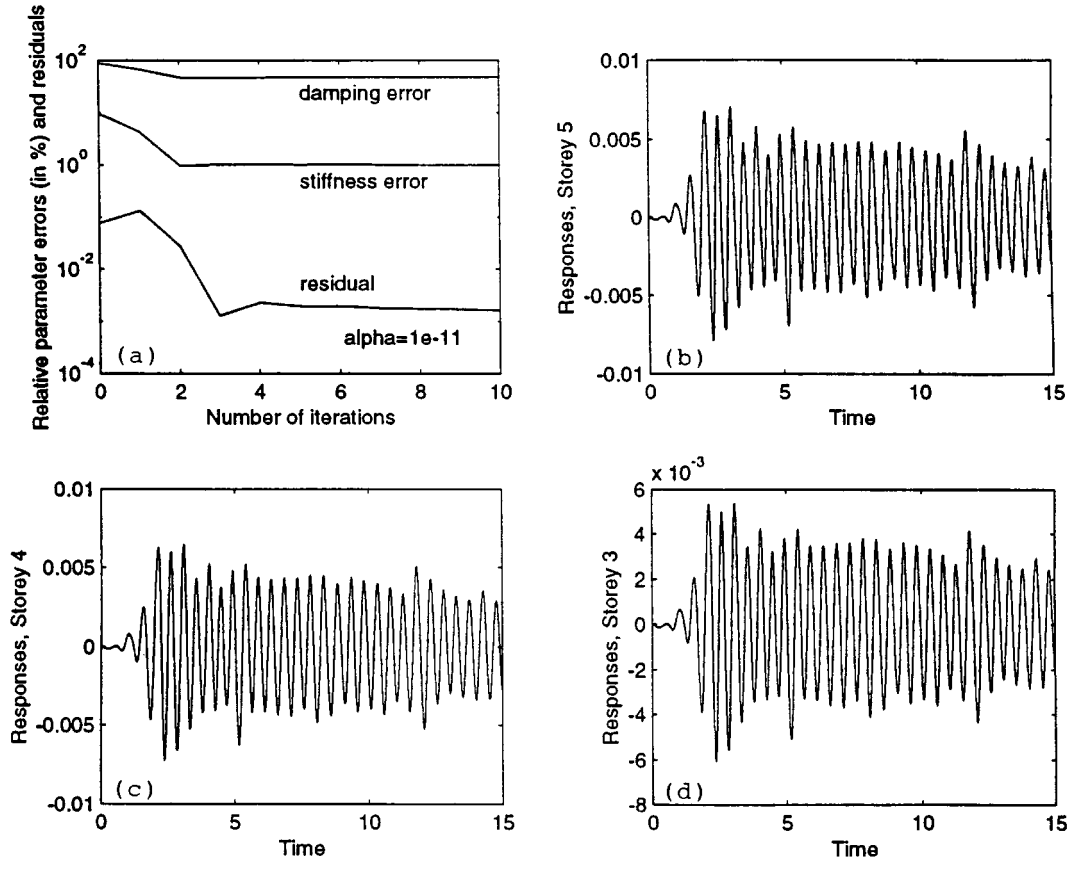

Figure 4. Coupled Van-der-Pol-Duffing oscillator with forcing function

Table IX. Coupled Van-der-Pol-Duffing oscillator with forcing function: errors in training vectors

	Five training vectors, $P^{(0)}$					Errors in $P^{(0)}$ in %					RMS error,%
	p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5	
a_1	6·0	4·5	4·7	5·4	4·0	− 20·0	10·0	6·0	− 8·0	20·0	14·14
b_1	2·2	1·8	2·3	3·0	2·5	− 10·0	10·0	− 15·0	− 50·0	− 25·0	26·65
c_1	4·8	5·3	4·0	4·7	5·2	4·0	− 6·0	20·0	6·0	− 4·0	10·04
d_1	2·3	1·8	2·4	1·9	1·8	− 15·0	10·0	− 20·0	5·0	10·0	13·04
e	1·3	0·85	1·2	1·1	1·5	− 30·0	15·0	− 20·0	− 10·0	− 50·0	28·72
a_2	4·4	3·7	4·2	3·0	4·3	− 10·0	7·5	− 5·0	25·0	− 7·5	13·13
b_2	3·0	2·7	2·2	2·8	2·3	− 20·0	− 8·0	12·0	− 12·0	8·0	12·77
c_2	5·2	4·8	5·3	4·3	4·9	− 4·0	4·0	− 6·0	14·0	2·0	7·32
d_2	3·3	2·8	3·2	3·6	2·8	− 10·0	6·7	− 6·7	− 20·0	6·7	11·25
ω	3·0	1·8	2·5	1·9	1·5	− 50·0	10·0	− 25·0	5·0	25·0	27·84

Table X. Coupled Van-der-Pol-Duffing oscillator with forcing function: errors in estimates, $\alpha = 10^{-9}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(8)}$		Exact param.
	Error %	Param.	Error %	Param.	
a_1	10·71	4·46	5·28	4·74	5·0
b_1	2·40	1·95	− 5·33	2·11	2·0
c_1	1·83	4·91	0·19	4·99	5·0
d_1	− 1·26	2·03	− 3·14	2·06	2·0
e	− 7·01	1·07	0·18	0·998	1·0
a_2	− 2·84	4·11	− 2·43	4·10	4·0
b_2	0·96	2·48	2·42	2·44	2·5
c_2	− 1·83	5·09	0·01	4·999	5·0
d_2	4·47	2·87	4·25	2·87	3·0
ω	− 2·50	2·05	− 0·0008	2·00002	2·0

time-windows are chosen for the identification, $t \in [0, 0{\cdot}6]$ and $t \in [1, 2{\cdot}2]$, and only measurements of $x(t)$ and $y(t)$ in these intervals are used for the identification. The data are concatenated to form the response vector r_j. The regularization parameter α is taken to be 10^{-9}.

Table X shows the primary estimate $\hat{p}(0)$ obtained after the five initial parameter vectors are used in the identification process along with the percentage error in each of the parameters. The number of training vectors genetically engendered at each iteration was $\bar{N} = 51$, $k > 0$. The results after eight iterations, starting with the primary estimate $\hat{p}(0)$, are also shown. Figure 5(a) shows the normalized error $\|\hat{p}^{(k)} - p^*\|/\|p^*\|$ in the parameter vector as a function of the iteration number k, along with the normalized response residual. The responses of the system (for $t \in [0, 5]$) using the exact parameter values and those estimated at the end of eight iterations are shown in Figures 5(b) and 5(c). At the scale plotted, no difference between the exact and the estimated responses can be seen.

Tables XI and XII show the results of the identification procedure, using a different set of five initial guess vectors for the same system as shown in equation (9), except that no forcing term on the right-hand side of equation (9b) is present. Measurements of $x(t)$ and $\dot{x}(t)$ are only used for the identification over the same two time windows. Figure 5(d) shows the normalized error in the parameter vector and that in the response residual. As before, the response using the estimated parameters and that using the exact parameter values are indistinguishable and are therefore not shown.

DISCUSSION AND CONCLUSIONS

In this paper we have used a different paradigm for the identification of structural systems than the one commonly used. Instead of using several inputs and several outputs to sequentially arrive at the parameters describing a dynamic system, we use a single input and sets of parameters values to match the measured or given output. Whereas previous methods start from a single parameter guess and sequentially update it to match the measured response, here we develop a sequential *set* of parameters and use them simultaneously to arrive at an estimate of the unknown parameter vector. Most of the computational time in our approach is spent on solving the forward problem instead of the inverse problem, another difference between standard approaches and the one used here. Since forward problems are considerably simpler to solve and have unique solutions (unlike inverse problems) the method appears to take good advantage of this.

The genetically engineered relevant training vectors are shown to improve the identification results dramatically by reducing the norm error in the parameter estimates significantly, especially in the absence of measurement noise. As opposed to prior work,[11,12] this significant increase in parameter estimation

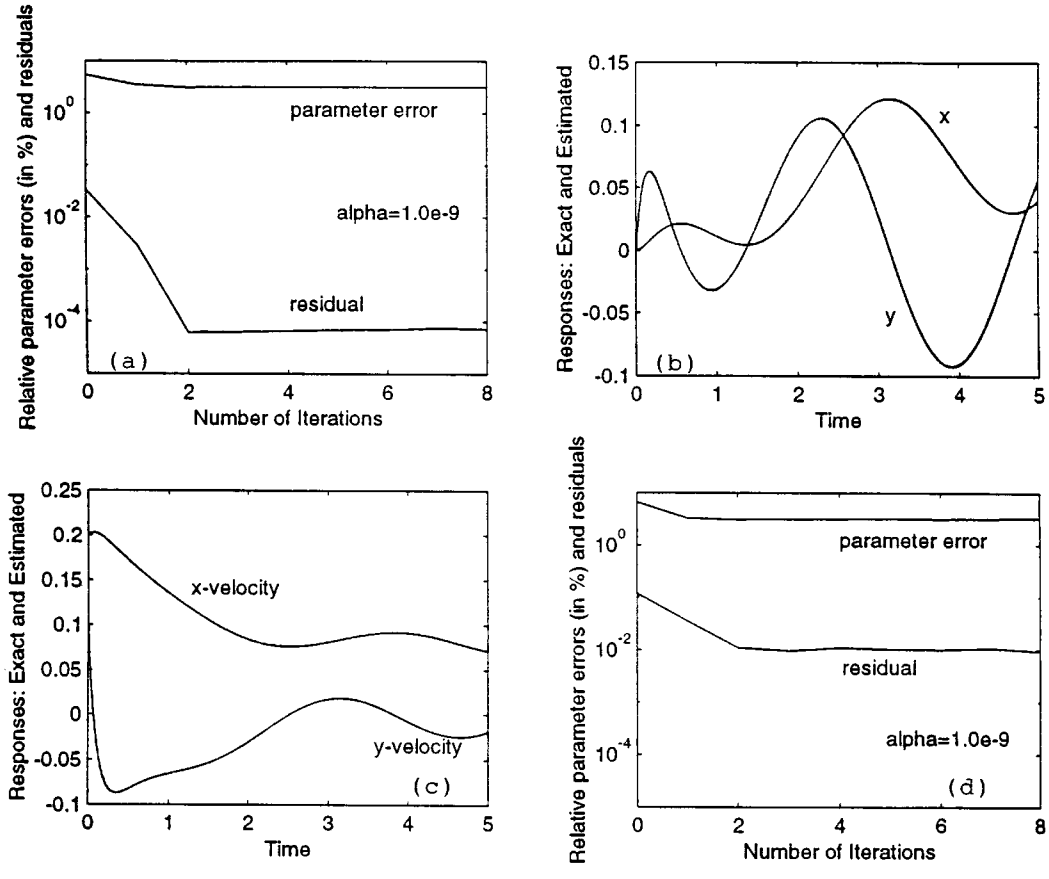

Figure 5. Coupled Van-der-Pol-Duffing oscillator without forcing function

Table XI. Coupled Van-der-Pol-Duffing oscillator without forcing function in equation (9b): errors in training vectors

	Five training vectors, $P^{(0)}$					Errors in $P^{(0)}$ in %					RMS error,%
	p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5	
a_1	5·5	4·5	4·7	5·4	5·0	10·0	− 10·0	− 6·0	8·0	0·0	7·75
b_1	2·2	1·8	2·3	1·9	2·5	10·0	− 10·0	15·0	− 5·0	25·0	14·66
c_1	4·8	5·3	5·0	4·7	5·2	− 4·0	6·0	0·0	− 6·0	4·0	4·56
d_1	2·3	1·8	2·4	1·9	1·8	15·0	− 10·0	20·0	− 5·0	− 10·0	13·04
e	1·3	0·85	1·2	1·1	1·0	30·0	− 15·0	20·0	10·0	0·0	18·03
a_2	4·4	3·7	4·2	4·0	4·3	10·0	− 7·5	5·0	0·0	7·5	6·89
b_2	2·5	2·7	2·2	2·8	2·3	0·0	8·0	− 12·0	12·0	− 8·0	9·12
c_2	5·2	4·8	5·3	4·3	4·9	4·0	− 4·0	6·0	− 14·0	− 2·0	7·32
d_2	3·3	2·8	3·2	3·0	2·8	10·0	− 6·7	6·7	0·0	− 6·7	6·83
ω	2·3	1·8	2·0	1·9	1·8	15·0	− 10·0	0·0	− 5·0	− 10·0	9·49

Table XII. Coupled Van-der-Pol-Duffing oscillator without forcing function in equation (9b): errors in estimates, $\alpha = 10^{-9}$

	Primary estimates, $\hat{p}^{(0)}$		Final estimates, $\hat{p}^{(8)}$		Exact param.
	Error %	Param.	Error %	Param.	
a_1	− 9·42	5·47	− 4·08	5·204	5·0
b_1	0·97	1·98	3·41	1·932	2·0
c_1	− 0·98	5·05	0·02	4·999	5·0
d_1	3·54	1·93	− 0·69	2·014	2·0
e	− 8·90	1·09	0·09	0·999	1·0
a_2	− 3·57	4·14	3·80	3·848	4·0
b_2	− 15·10	2·88	− 5·22	2·630	2·5
c_2	6·35	4·68	− 1·03	5·051	5·0
d_2	− 2·29	3·07	− 6·47	3·194	3·0
ω	− 0·99	2·02	− 0·31	2·006	2·0

accuracy can be attributed to the combined effect of the following four factors: (1) the generation of an entirely new set of training vectors at each iteration; (2) the generation of each set of training vectors (using random perturbations of the estimated parameter) through a selection process which accepts only those vectors whose response residuals are small; the variance of the random perturbations is adaptively changed with each iteration; (3) the use of small values of the regularization parameter, α; and (4) the development of an adaptive scheme which is driven by the information acquired from the response residuals. We demonstrate that this new memory matrix approach proposed herein is thus a powerful approach for identification of the physical parameters of a system even when these parameters may differ by several orders of magnitude. The approach provides remarkably accurate results while using a very small amount of data when compared with recursive-prediction-error methods or other hill-climbing methods. The latter methods require long data windows (and large amounts of data) and even with such long windows often yield poor identification results in such cases, see Reference 12. It is this feature of our methodology which may make it promising for real-time, on-line, quick identification and health monitoring of complex structural and mechanical systems.

REFERENCES

1. R. E. Kalaba and K. Spingarn, *Control, Identification and Input Optimization*, Plenum, New York, 1982.
2. G. Rodriguez (ed.), *Proc. Workshop on Identification and Control of Flexible Structures*, Vols. I and II, JPL Publication 85-29, April 1985.
3. J. Hopfield and D. Tank, 'Computing with neural circuits: a model', *Science* **233**, 625–633 (1986).
4. T. Kohonen, *Self-organization and Associative Memory*, Springer, New York, 1988.
5. L. Ljung, *System Identification: Theory for the User*, McGraw-Hill, New York, 1988.
6. B. Kosko, 'Bidirectional associative memories', *IEEE Trans. Systems Man Cybernet.* **SMC-18**, 49–60 (1988).
7. F. Udwadia, J. Garba and A. Ghodsi, 'Parameter identification problems in structural and geotechnical engineering', *J. Engng. Mech.* **110**, 1409–1432 (1984).
8. F. Udwadia and D. Sharma, 'Some uniqueness problems in the identification of building structures', *SIAM J. Appl. Math.* **34**, 104–118 (1978).
9. A. Sage and J. Melsa, *Estimation Theory with Applications to Communication and Control*, McGraw-Hill, New York, 1971.
10. D. Rehak, C. Thewalt and L. Doo, 'Neural network applications in structural mechanics computation', *J. Comput. Utilization*, 20–31 (1989).
11. R. Kalaba and F. Udwadia, 'An adaptive learning approach to the identification of structural and mechanical systems', *Comput. Math. Appl.* **22**(1), 67–75 (1991).
12. R. Kalaba and F. Udwadia, 'An associative memory approach to the rapid identification of nonlinear structural and mechanical systems', *J. Optim. Theory Appl.* **76**(2), 207–223 (1993).
13. J. C. Lin and D. M. Durand, 'Weighted linear associative memory approach to nonlinear parameter identification', *J. Optim. Theory Appl.* **90**(1), 139–159 (1996).

